Potential outcomes and randomized experiments

Frank Venmans

University of Mons

frank.venmans@umons.ac.be

November 28, 2016

Overview

References

- Angrist D., Pischke J.-S. (2009) Mostly harmless econometrics. Princeton University Press, Princeton.
- Verbeek M., (2012) Modern guide to econometrics. 4th edition. Wiley.
- Wooldridge J.M. (2010) The econometric analysis of Cross Section and Panel Data. 2nd edition, MIT Press, Cambridge.

Cause, effect and causal relationship

According to John Stuart Mill, causal relationship satisfies three criteria:

1) the "cause" precedes the "effect" in time (temporal precedence),
2) the "cause" and the "effect" are related (covariation), and
3) there are no plausible alternative explanations for the observed covariation (nonspuriousness).
(Shadish, Cook, Campbell 2002)

The filosophical discussion about causality and different types of causality, dating back at least to Aristotle, is beyond the scope of this course.

Causal effect of a treatment

We want to know what's the causal effect of a "treatment":

Units	**Treatment**	**Outcome**
Plots	Fertilizer	Plant growth
Children	Extra year of schooling	Wage
Adults	Training program	Employment
Companies	Minimum wage	Employment
Class	Class size	School attainment

Potential outcomes (Neyman 1923)

Treatement

D_i: Treatment indicator for unit i

$$D_i \begin{cases} 1 & \text{if unit i received the treatment} \\ 0 & \text{otherwise.} \end{cases}$$

Observed outcome

Y_i: Observed outcome variable for unit i

Potential outcome

$$Y_{di} \begin{cases} Y_{1i} & \text{Potential outcome for unit i with treatment} \\ Y_{0i} & \text{Potential outcome for unit i without treatment} \end{cases}$$

The unobserved potential outcome is is called the counterfactual

Rubin's causal model (1974)

Causal effect

$$\alpha_i = (Y_{1i} - Y_{0i})$$

There may be a different treatment effect for different units

Stable Unit Treatment Value Assumption (SUTVA)

1. Potential oucomes for unit i are unaffected by treatment assignment for unit j. Non-interference among units: treatment T_i only impacts unit i
   - ex: Effect of flu vaccine on hospitalization
   - ex: General equilibirium effects
2. The treatments for all units are comparable (no variation in treatment)
   - ex: training program differs in different cities or periods

ATE and ATET

The effect of a job training program on employment may be different for a low-skilled person compared to a high-skilled person.
Therefore, the effect on the whole population (ATE) will be different than the effect on a subset of population that finally takes the program(ATET).

Average Treatment Effect (ATE)

$$\alpha_{ATE} = E[Y_1 - Y_0]$$

Average Treatment Effect of the Treated (ATET or ATT)

$$\alpha_{ATET} = E[Y_1 - Y_0|D = 1]$$

ATE and ATET

Imagine a population of 4 unemployed people. 2 low-skilled persons take the training program, whereas 2 high-skilled persons do not take the program.
Y is income 1 year after program.
Numbers in red are unobserved.

i	Y_{1i}	Y_{0i}	Y_i	D_i	α_i
1	5	4	5	1	1
2	4	2	4	1	2
3	6	7	7	0	-1
4	8	7	7	0	1

ATE= (1+2-1+1)/4=0.75
ATET= (1+2)/2=1.5
Note that comparing the observed treated with the observed untreated does not yield a correct causal effect

The effect of being hospitalized on health

Going to a hospital has beneficial effect (the medical treatment) but also potential negative effects (getting infected by others, poor diagnosis).

Group	Sample Size	Mean Health Status	Std. Error
Hospital	7,774	3.21	0.014
No hospital	90,049	3.93	0.003

- NHIS 2005 survey
- Health status: 1=poor; 5=excellent health
- Does going to a hospital lead to lower health status?
- Probably, people who went to the hospital were already less healthy before visiting the hospital....

Selection bias

Comparing mean outcomes between treated and untreated

$$= E[Y_1|D=1] - E[Y_0|D=0]$$

$$= \underbrace{E[Y_1 - Y_0|D=1]}_{ATET} + \underbrace{\{E[Y_0|D=1] - E[Y_0|D=0]\}}_{Bias}$$

The health of the hospitalized people in case they would not have gone to the hospital ($E[Y_0|D=1]$) is probably not the same as the health of those people who did not go to the hospital ($E[Y_0|D=0]$), resulting in a bias.

Bias can be positive or negative. Therefore, comparing treated and untreated does not mean anything.

Correlation is not a necessary nor a sufficient condition for causality!

Solutions to Selection bias

- Randomized experiments
- Selection on observables (regression, matching, weighting)
- Selection on unobservables (difference-in-difference, instrumental variables, regression discontinuity design)

Randomized experiments

Fisher was first to propose randomization in experiments.
A lady asserts that she can taste whether the milk or the tea was added first to a cup.
Randomly assign eight cups to 2 conditions:

- Tea first (control)
- Milk first (treatment)

Ask the lady to discriminate between tea first and milk first

Fisher's experiment(1935)

Why randomization and not ceteris paribus conditions?

"It is no sufficient remedy to insist that "all of the cups must be exactly alike" in every respect except that to be tested. For this is a totally impossible requirement in our example, and equally in all other forms of experiment. In practice it is probable that the cups will differ perceptibly in the thickness or smoothness of their material, that the quantities of milk added to the infusions may change between pouring the first and the last cup, and that the temperature also at which the tea is tasted will change during the course of the experiment..." (p 18)

It is impossible to keep observable and unobservable confounders constant. Therefore, render confounders impotent by randomization.

Randomized experiments

- How would you measure the effect of hospitalization for those that take the decision to go to the hospital, using a randomized experiment? Discuss with your neighbour.
- Randomly send back people that want to be cured in the hospital and compare their health with people that were accepted.
- As shown in this example, in many cases, truly random experiments are impossible in the social sciences.

randomized experiments

identification assumption

$(Y_1, Y_0) \perp D$ (Outcomes are independent from assignment, or random assignment)

identification result

Random assignment implies $E[Y_0|D=1] = E[Y_0|D=0]$

$$\implies \alpha_{ATET} = E[Y_1 - Y_0|D=1] = E[Y_1|D=1] - E[Y_0|D=1]$$

$$= E[Y_1|D=1] - E[Y_0|D=0]$$

Random assignment also implies $E[Y_1|D=1] = E[Y_1|D=0] \implies$

$$\alpha_{ATE} = E[Y_1 - Y_0] = E[Y_1] - E[Y_0] = E[Y_1|D=1] - E[Y_0|D=0]$$

Difference in means$=\alpha_{ATET} = \alpha_{ATE}$

Testing in a large sample: Two sample T-test

- For any probability distribution of Y, the estimated mean $\bar{Y}$ will converge to a normal distribution when the sample size increases (consequence of central limit theorem).
- It can therefore be shown that $t = \frac{\hat{\alpha}}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_0^2}{N_0}}}$ follows a t-distribution
- We reject the null hypothesis $H_0 : \alpha = 0$ at the asymptotic 0.05 significance level if $|t| > 1.96$
- For small samples where Y does not follow a normal distribution: Fisher's Exact Test

Covariate balance

- Randomization balances not only outcome variable Y but also
  - observed characteristics of treatment and control group
  - unobserved characteristics of treatment and control group
- You can check random assignment with a balance test
  - test if pre-treatment variable X is the same for treated and control group.
  - Example t-test: $H_0 : E(X|D = 0) = E(X|D = 1)$

Threats to internal validity

Can we estimate treatement effect for our particular sample?
Fails when there are differences between treated an control (other than the treatment itself) that affect the outcome and that we cannot cotrol for.
Most common problems

- Failure of randomization (effect of confounders)
- Non-compliance with experimental protocol (treatment is not the same for everybody)
- Attrition (drop out)

Threats to external validity

Can we extapolate our estimates to other populations (other than our sample)?
Fails when outside the sample, the treatment has a different effect
Most common problems

- Non representative sample
- Non representative program
  - Treatment differs in actual implementations
  - Scale effects
  - Actual implementations are not randomized (nor full scale)